International Journal of Solids and Structures 46 (2009) 2463–2476

Contents lists available at ScienceDirect

International Journal of Solids and Structures

journal homepage: www.elsevier.com/locate/ijsolstr

Thermoelastic instability in friction clutches and brakes – Transient modal analysis revealing mechanisms of excitation of unstable modes

Przemyslaw Zagrodzki *

Friction Holdings LLC, Technical Center, 731 Tech Drive, Crawfordsville, IN 47933, USA

ARTICLE INFO

Article history:
Received 29 September 2008
Received in revised form 30 December 2008
Available online 11 February 2009

Keywords:
Sliding contact
Friction clutch/brake
Thermoelastic instability
Transient
Finite element method
Modal analysis
Modal superposition

ABSTRACT

Sliding systems with frictional heating exhibit thermoelastic instability (TEI) when the sliding speed exceeds the critical value. TEI can lead to hot spots on contact surfaces and is generally of great practical importance in friction brakes and clutches. The phenomenon is well defined in terms of the theory of stability with a classic perturbation approach being commonly used. While the perturbation analysis determines the stability limit, recent interest extends further towards exploration of the unstable behavior. This is motivated by practical reasons, namely by the fact that many common friction brakes and clutches operate instantaneously at speeds exceeding the critical speed for TEI, i.e. in the unstable regime. In order to determine a transient solution, possible mechanisms of excitation of unstable modes of different nature need to be accurately defined and quantified. These mechanisms are normally not considered in stability analysis of the steady-state where an initial perturbation of the thermoelastic field is assumed. In many realistic situations, however, there is no indication of the existence of meaningful initial temperature variation. Lack of full understanding of these mechanisms has perhaps limited broader industrial applications of recent theoretical advances in TEI. In this paper a method of solving the transient thermoelastic process in frictional systems using finite element spatial discretization and modal superposition is presented. Then mechanisms that excite the unstable thermoelastic modes other than the initial perturbation of temperature are studied. The role of the background process (corresponding to nominal applied loads) in the excitation is shown in a clear form and illustrated by practical examples for automotive friction clutches. It is demonstrated, in particular, that while for some geometries and configurations of the sliding system the imperfections determine the excitation of unstable modes, with other configurations strong excitation occurs even in the absence of imperfections.

1. Introduction

The frictional heat flux generated at the contact interface of a sliding system is proportional to the contact pressure. Therefore, if some non-uniformity in pressure distribution across the surface occurs, the areas where the pressure is higher experience a higher temperature increase. This in turn causes greater local thermal expansion and thereby leads to further local pressure increase. Thus, the sliding system with frictional heating inherits a mechanism that tends to magnify contact pressure non-uniformity. This phenomenon, identified by Barber (1969), is known as frictionally excited thermoelastic instability (TEI). TEI can lead to local contact concentrations that manifest themselves by hot spots on the friction surfaces (Barber, 1969; Anderson and Knapp, 1990; Zagrodzki and Truncone, 2003; Kao et al., 2000; Lee and Brooks, 2003) and it has profound practical consequences in clutches and brakes as well as other sliding systems with frictional heating.

The phenomenon is well defined in terms of the theory of stability, with the perturbation approach being commonly used (e.g. Dow and Burton, 1972; Lee and Barber, 1993a; Du et al., 1997; Yi et al., 2000; Decuzzi et al., 2001; Krempaszky and Lippmann, 2005). While the classic perturbation analysis determines stability limits for the system, recent interest is directed towards exploration of the unstable behavior (Zagrodzki, 1990; Kao et al., 2000; Zagrodzki et al., 2001; Al-Shabibi and Barber, 2002; Afferrante et al., 2003; Choi and Lee, 2004). This is motivated by practical considerations, namely by the fact that many common friction brakes and clutches operate instantaneously at speeds exceeding the critical speed for TEI, i.e. in the unstable regime (Yi et al., 2000; Zagrodzki and Truncone, 2003). Exploration of the transient thermoelastic process to determine instantaneous temperatures, thermal deformations and stresses occurring in these conditions is therefore of great interest.

A linear model of the thermoelastic process in a brake or a clutch can usually be adopted for the stage of operation with full contact. On the other hand, it is known that hot spots produced even during short-term slip can be accompanied by local separation of the surfaces (Zagrodzki and Truncone, 2003). This is certainly an

* Tel.: +1 765 359 2567; fax: +1 765 359 2566.
E-mail address: pzagrodzki@frictionholdings.com

0020-7683/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijsolstr.2009.01.033

undesirable situation that should be prevented and control of thermoelastic behavior is aimed at preservation of full contact. Thus, a linear model, without separation, in spite of its limitations, is useful in this type of practical application. We can distinguish two components of the thermoelastic process: the background process, which corresponds to contact pressure resulting from the applied external loads (not necessarily uniform or constant), and the perturbed process, resulting from the initial condition, superposed on the background process. Stability of a linear system does not depend on the background process and therefore this process is typically not included in the stability analysis (Dow and Burton, 1972; Lee and Barber, 1993a; Du et al., 1997; Yi et al., 2000; Decuzzi et al., 2001; Krempaszky and Lippmann, 2005). Also, in many known studies of transient behavior, e.g. Al-Shabibi and Barber (2002), Afferrante et al. (2003) and Voldrich (2007), an initial perturbation of temperature is assumed and the background process is not considered. Initial temperature variation seems to be unquestionable in some practical situations, for instance as a remnant of the preceding clutch or brake application. In many other important practical instances, however, hot spots were observed without any indication of a meaningful initial temperature variation. For example, severe hot spots were detected after a single application of a brake or a clutch by Anderson and Knapp (1990) and Zagrodzki and Truncone (2003). These observations indicate that factors other than the initial perturbation of temperature can excite the unstable process. The background process seems to play major role by producing temperature variation which contains components consistent with the unstable modes.

Understanding of these processes is crucial to the solution of important industrial problems. An example is the resolution of a hot spotting problem in an automotive transmission clutch by means of re-arrangement of clutch components, conceived as a result of identification of the mechanism exciting the unstable modes (Zagrodzki and Zhao, 2008).

In the paper, modal analysis is used in the solution of the thermoelastic problem in the sliding frictional system. More specifically, spatial discretization of the elastic and thermal problems is performed using the finite element method. In the thermal model the relative motion of system components is accounted for (in-plane sliding). Then the two problems are coupled by interfacial boundary conditions with frictional heating and a semi-discrete, time-dependent problem is obtained. The general solution of this problem is presented in analytical form. The semi-discrete problem is then reduced to an eigenvalue problem with eigenvalues representing the stability parameter, an approach similar to that used in Yi et al. (2000) and Al-Shabibi and Barber (2002). Next, by transforming the problem to modal variables, the transient solution is expressed by modal superposition; this solution was presented by Zagrodzki (2003), and used by Li and Barber (2004) and Li and Barber (2008). This comprehensive formulation clearly exhibits contributions from the initial conditions and from the nonhomogeneous part representing the background process and helps to elucidate the relations between them. In the background process, geometric imperfections are considered but also important practical cases are identified where the unstable process is excited in the absence of such imperfections.

The background process was included in recent study by Li and Barber (2008), who presented the solution for an axisymmetric problem, similar to one of the cases discussed in this paper. Note that the background process is also typically included in finite element simulations in time domain (Zagrodzki, 1990; Kao et al., 2000; Zagrodzki et al., 2001; Zagrodzki and Truncone, 2003; Choi and Lee, 2004; Zagrodzki and Zhao, 2008). This method, robust in tackling non-linear problems, does not offer the clarity of modal superposition and therefore provides only limited insight into the contributing mechanisms.

This paper concentrates on the method of transient modal analysis and the elucidation of the mechanisms exciting unstable modes. For convenience, the analysis is performed for constant sliding speed. Substantial contributions to the solution of problems with variable speed using modal analysis have been made by Al-Shabibi and Barber (2002) and then further enhanced by Li and Barber (2004) and Li and Barber (2008); the latter approach was proven to be particularly robust and it can be applied to the models used in this study.

2. Model and discretization of thermoelastic problem in the sliding system

The overall formulation of the problem and the method of solution used in this study were presented by Zagrodzki (2003). The general case of a sliding system composed of two or more solids with frictional heat generation at contact interfaces is considered. For illustration purposes, without loss of generality, a two-dimensional system comprising three components with two sliding interfaces shown in Fig. 1 is used throughout the most of this paper. This model represents a multidisk clutch or a brake with the central layer corresponding to the metal disk and the two external layers representing the friction material (Lee and Barber, 1993a; Zagrodzki and Truncone, 2003). The mechanical part of the problem is treated as a quasistatic elastic stress problem subjected to external loads and thermal strains. At the sliding interfaces, continuity of normal displacements is assumed and relative tangential displacements are allowed. The thermal part is described by a transient heat transfer problem. At the sliding interfaces temperature continuity is assumed and a frictional heat input is applied as a function of local contact pressure and speed. As a result, a fully coupled, time-dependent thermoelastic problem is obtained.

The stage of operation of the sliding system during which full contact occurs is considered. This assumption along with an assumption of linear material properties leads to a linear problem, which is treatable by modal analysis. Shear tractions, that have been shown to play only a minor role in most practical thermoelastic contact problems (Lee and Barber, 1993b), are neglected. This last assumption simplifies the problem without loss of generality.

The finite element method is used for spatial discretization of both the elastic and the thermal parts of the problem.

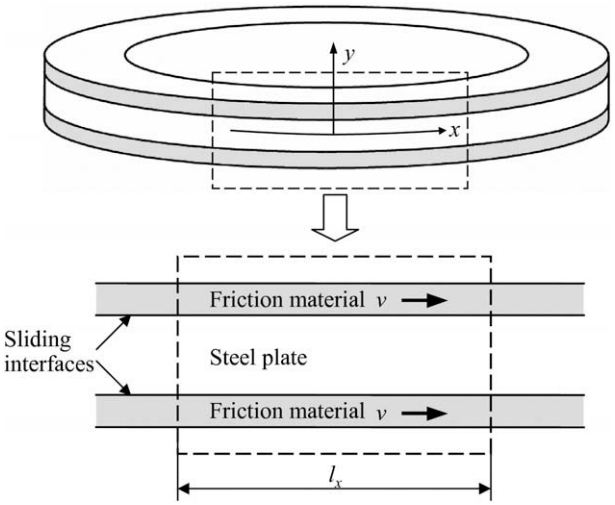

Fig. 1. Model geometry, in-plane sliding.

2.1. Interfacial pressure due to temperature and external loads

Detailed derivations for the finite element discretization of the elastic problem are included in Appendix A. The contact pressure at the sliding interfaces is represented by equivalent concentrated nodal forces which can be expressed as a linear function of temperature and applied external loads

$$\boldsymbol{P} = \boldsymbol{AT} + \boldsymbol{BF}_a, \tag{1}$$

where $\boldsymbol{P}$ is a column vector of nodal forces at all interfaces in the sliding system, $\boldsymbol{T}$ is the vector of nodal temperatures and $\boldsymbol{F}_a$ is the vector of external forces applied to the system. $\boldsymbol{A}$ and $\boldsymbol{B}$ are rectangular matrices. The component $\boldsymbol{BF}_a$ also includes the effects of possible geometric imperfections of the contact surfaces as described in Appendix A.

2.2. Heat transfer

The unsteady heat transfer problem is formulated in Lagrangian–Eulerian coordinates. Spatial discretization of this problem by finite elements is presented in Appendix B. It yields a semi-discrete system in the form

$$\boldsymbol{C}\dot{\boldsymbol{T}} + (\boldsymbol{L_D} + v_c\boldsymbol{L_C})\boldsymbol{T} + \boldsymbol{Q} = \boldsymbol{0}, \tag{2}$$

where $\boldsymbol{C}$ is the thermal capacitance matrix, $\boldsymbol{L_D}$ is the thermal conductivity matrix, $v_c\boldsymbol{L_C}$ is the convection matrix, v_c is the speed of convection and $\boldsymbol{Q}$ is the vector of nodal heat sources at the sliding interfaces. These heat sources represent frictional heating and are given by

$$\boldsymbol{Q} = f v_s \boldsymbol{P}, \tag{3}$$

where f is the friction coefficient and v_s is the sliding speed. The speed of convection v_c and the sliding speed v_s can be identical (as in the model in Fig. 1 with in-plane sliding and the coordinate system fixed to one of the components) but in general the two speeds are distinct.

2.3. Thermoelastic problem with frictional heating

Substituting Eqs. (1) and (3) into Eq. (2) leads to the following equation

$$\boldsymbol{C}\dot{\boldsymbol{T}} + (\boldsymbol{L_D} + v_c\boldsymbol{L_C} + f v_s\boldsymbol{A})\boldsymbol{T} = -f v_s \boldsymbol{BF_a}. \tag{4}$$

Eq. (4) represents the transient thermoelastic problem. Note that by using the above substitution the displacement degrees of freedom are eliminated from the problem and only temperatures are retained.

As noted in Appendix B, matrices $\boldsymbol{C}$ and $\boldsymbol{L_C}$ are non-symmetric while matrix $\boldsymbol{L_D}$ is symmetric. Matrix $\boldsymbol{A}$ is non-symmetric.

Eq. (2) describes diffusive-convective heat transport and, by arguments based on physics, has only stable solutions. The presence of the term $fv_s\boldsymbol{A}$ on the left side of Eq. (4) modifies the system's properties so that it can become unstable. For this to happen, the product fv_s has to be sufficiently high. If the friction coefficient f is considered a constant parameter, the sliding speed v_s determines stability of the system. The threshold speed above which the system is unstable is referred to as the critical speed (Dow and Burton, 1972).

Eq. (4) represents a linear problem and its nonhomogeneous part is irrelevant to stability. Therefore, in stability studies (e.g. Du et al., 1997; Yi et al., 2000) only the homogeneous problem, with the right side vector set to zero, was considered. These studies determined stability characteristics described by the growth rate of a perturbation. They defined the critical sliding speeds, and the associated thermoelastic modes but did not address the general transient solution. As mentioned, many sliding systems, like transmission clutches or brakes, operate instantaneously at high speeds that significantly exceed the critical speed. In this case the transient solution and ultimate variation of temperature/pressure after some period of slip in the super-critical regime is of great interest. Eq. (4) can be written in the form

$$\dot{\boldsymbol{T}} - \boldsymbol{GT} = \boldsymbol{SF_a}(t), \tag{5}$$

where

$$\begin{aligned} \boldsymbol{G} &= -\boldsymbol{C^{-1}}(\boldsymbol{L_D} + v_c\boldsymbol{L_C} + f v_s\boldsymbol{A}), \\ \boldsymbol{S} &= -f v_s \boldsymbol{C^{-1}B}. \end{aligned} \tag{6}$$

Here $\boldsymbol{G}$ is the system matrix. The properties of $\boldsymbol{G}$ determine, in particular, the stability of the system.

3. General solution of thermoelastic equation

We assume that speeds v_s, v_c = const. and, consequently, $\boldsymbol{G},\boldsymbol{S}$ = const. The general solution of the ordinary linear differential equation, Eq. (5), then has the form (Coddington and Carlson, 1997)

$$\boldsymbol{T}(t) = \exp(\boldsymbol{G}t)\boldsymbol{T}(0) + \int_0^t \exp[\boldsymbol{G}(t-\tau)]\boldsymbol{SF_a}(\tau)\,d\tau, \tag{7}$$

where the matrix exponential function is defined in the known way

$$\exp(\boldsymbol{G}t) = \boldsymbol{I} + \frac{\boldsymbol{G}t}{1!} + \frac{\boldsymbol{G}^2t^2}{2!} + \frac{\boldsymbol{G}^3t^3}{3!} + \cdots$$

and $\boldsymbol{I}$ is the identity matrix of the same dimension as $\boldsymbol{G}$. The first term on the right side of Eq. (7) represents the part of the solution corresponding to the homogeneous equation, i.e. to Eq. (5) with $\boldsymbol{F}_a(\mathrm{t}) \equiv \boldsymbol{0}$. In other words, it represents the solution corresponding to the initial condition $\boldsymbol{T}(0)$ and it is obvious that $\boldsymbol{T}(0) = \boldsymbol{0}$ implies a trivial solution within any finite time interval, even if the system is unstable. The second term on the right side of Eq. (7), containing the convolution integral, is the solution corresponding to the nonhomogeneous part of Eq. (5). To conveniently interpret the behavior of the thermoelastic system, we express solution (7) in terms of the eigenvectors of the matrix $\boldsymbol{G}$.

$\boldsymbol{G}$ is a real, non-symmetric matrix. If the matrix has linearly independent eigenvectors (which is also possible with multiple eigenvalues), matrix $\exp(\boldsymbol{G}t)$, known as the transition matrix, can be expressed as (Deif, 1982)

$$\exp(\boldsymbol{G}t) = \boldsymbol{V}\exp(\boldsymbol{\Lambda}t)\boldsymbol{V}^{-1}, \tag{8}$$

where

$\boldsymbol{V}$ is the matrix of eigenvectors of $\boldsymbol{G}$; $\boldsymbol{\Lambda}$ is the (diagonal) matrix of eigenvalues of $\boldsymbol{G}$. By substituting Eq. (8) into Eq. (7) we obtain

$$\boldsymbol{T}(t) = \boldsymbol{V}\exp(\boldsymbol{\Lambda}t)\boldsymbol{V^{-1}T}(0) + \int_0^t \boldsymbol{V}\exp[\boldsymbol{\Lambda}(t-\tau)]\boldsymbol{V^{-1}SF_a}(\tau)\,d\tau. \tag{9}$$

In the case where the right side of Eq. (5) is constant, $\boldsymbol{F}_a(\mathrm{t}) \equiv$ const., Eq. (9) takes the form

$$\boldsymbol{T}(t) = \boldsymbol{V}\exp(\boldsymbol{\Lambda}t)\boldsymbol{V^{-1}T}(0) + \boldsymbol{V}\int_0^t \exp[\boldsymbol{\Lambda}(t-\tau)]\,d\tau\boldsymbol{V^{-1}F} \tag{10}$$

or

$$\boldsymbol{T}(t) = \boldsymbol{V}\exp(\boldsymbol{\Lambda}t)\left[\boldsymbol{V^{-1}T}(0) + \int_0^t \exp(-\boldsymbol{\Lambda}\tau)\,d\tau\cdot\boldsymbol{V^{-1}F}\right], \tag{11}$$

where

$$\boldsymbol{F} = \boldsymbol{SF}_a. \tag{12}$$

3.1. Transformation to modal variables

The modal variables are defined as follows:

$$\widetilde{\boldsymbol{T}}(t) = \boldsymbol{V}^{-1}\boldsymbol{T}(t). \tag{13}$$

The solution in modal variables has the form

$$\widetilde{\boldsymbol{T}}(t) = \exp(\boldsymbol{\Lambda}t)\left[\widetilde{\boldsymbol{T}}(0) + \int_0^t \exp(-\boldsymbol{\Lambda}\tau)\,d\tau\widetilde{\boldsymbol{F}}\right], \tag{14}$$

where

$$\widetilde{\boldsymbol{T}}(0) = \boldsymbol{V}^{-1}\boldsymbol{T}(0) \tag{15}$$

is the initial condition in modal variables and

$$\widetilde{\boldsymbol{F}} = \boldsymbol{V}^{-1}\boldsymbol{F} \tag{16}$$

is the vector of applied loads transformed to nodal variables.

Note that since $\boldsymbol{\Lambda}$ is a diagonal matrix, so is the matrix $\exp(\boldsymbol{\Lambda}t) = \mathrm{diag}\,[\exp(\lambda_1 t),\ldots,\exp(\lambda_N t)]$ and, consequently, the solution in modal variables given by Eq. (14) is a set of uncoupled equations. Here N is the dimension of the matrix. Eq. (14) can be written in scalar form as

$$\widetilde{T}_i(t) = \exp(\lambda_i t)\left[\widetilde{T}_{0i} + \int_0^t \exp(-\lambda_i\tau)\,d\tau\widetilde{F}_i\right]; \quad i = 1,\ldots,N, \tag{17}$$

where $\widetilde{T}_{0i} = \widetilde{T}_i(0)$. The integrals in Eq. (17) can easily be evaluated as

$$\int_0^t \exp(-\lambda_i\tau)\,d\tau = \begin{cases} \frac{1}{\lambda_i}[1-\exp(-\lambda_i t)] \text{ when } \lambda_i \neq 0 \\ t \text{ when } \lambda_i = 0 \end{cases}; \quad i = 1,\ldots,N. \tag{18}$$

3.2. Interpretation of the solution of thermoelastic problem

Eq. (17) represents the general solution of the thermoelastic problem, that is the solution corresponding to both the initial condition $\widetilde{T}_{0i}$ (homogeneous problem), and the external load $\widetilde{F}_i$ (nonhomogeneous problem). If the real part of any ith eigenvalue in Eq. (17) is positive, $\mathrm{Re}(\lambda_i) > 0$, the solution is unstable. The magnitude of the real part is a measure of stability and it is referred to as the growth rate or the degree of instability. Since we are interested in the transient evolution of the solution during a limited time interval, the degree of instability is an important characteristic.

Either the initial temperature $\widetilde{T}_{0i}$ or the external load $\widetilde{F}_i$ can produce a contributor from the ith mode, whether it is stable or unstable. Unlike the initial condition, the part corresponding to $\widetilde{F}_i$ is always zero at time $t = 0$ according to Eq. (18). If the mode is unstable ($\mathrm{Re}(\lambda_i) > 0$), the component $\exp(-\lambda_i t)$ in Eq. (18) decays asymptotically and after some transitional phase the load $\widetilde{F}_i$ acts analogously to the initial perturbation. The higher the degree of instability, the shorter is this transitional phase.

$\widetilde{T}_{0i} \neq 0$ certainly means that the initial temperature contains component corresponding to the ith eigenmode. Similarly, $\widetilde{F}_i \neq 0$ denotes that the external load $\boldsymbol{F}$ has a contribution compatible with ith eigenmode or, in other words, is non-orthogonal with respect to this mode, which can clearly be seen from Eq. (16) re-written in the form $\boldsymbol{F} = \boldsymbol{V}\widetilde{\boldsymbol{F}}$.

3.3. Solution in nodal temperatures expressed by modal variables

We consider the general solution in the original (physical) variables $\boldsymbol{T}(t)$, Eq. (11). Using Eqs. (17) and (18), a contribution from the ith eigenmode to this solution can be expressed as

$$\begin{aligned} \boldsymbol{T}_i(t) &= \boldsymbol{V}_i \exp(\lambda_i t)\cdot\left\{\widetilde{T}_{0i} + \left[\frac{1}{\lambda_i}[1-\exp(-\lambda_i t)]\right]\cdot\widetilde{F}_i\right\} \text{ when } \lambda_i \neq 0, \\ \boldsymbol{T}_i(t) &= \boldsymbol{V}_i\left(\widetilde{T}_{0i} + \widetilde{F}_i\cdot t\right) \text{ when } \lambda_i = 0. \end{aligned} \tag{19}$$

The eigenvectors $\boldsymbol{V}_i$ and the corresponding eigenvalues λ_i are generally complex since the system matrix $\boldsymbol{G}$, Eq. (6), is non-symmetric. Consequently, $\boldsymbol{T}_i(t)$ in Eq. (19) is generally complex. However, since $\boldsymbol{G}$ is real, complex eigenvectors occur as pairs of conjugate vectors $\boldsymbol{V}_i$ and $\bar{\boldsymbol{V}}_i$ and, similarly, there are conjugate pairs of corresponding eigenvalues λ_i and $\bar{\lambda}_i$. It is easy to verify that the initial conditions $\widetilde{T}_{0i}$ corresponding to any pair of conjugate eigenvectors, are also mutually conjugate; the same applies to $\widetilde{F}_i$ and, consequently, to $\boldsymbol{T}_i(t)$. Therefore, instead of complex vectors $\boldsymbol{T}_i(t)$ we will consider real vectors $\boldsymbol{T}_{\prime i}(t) = \boldsymbol{T}_i + \bar{\boldsymbol{T}}_i$ corresponding to a pair of conjugate eigenmodes. Denoting $\boldsymbol{V}_i^{\mathrm{re}} = \mathrm{Re}(\boldsymbol{V}_i)$, $\boldsymbol{V}_i^{\mathrm{im}} = \mathrm{Im}(\boldsymbol{V}_i)$ and using a similar notation for the real and imaginary parts of λ_i, $\widetilde{T}_{0i}$ and $\widetilde{F}_i$, the following equations were obtained after necessary manipulations:

- For the homogeneous part of the solution (corresponding to the initial condition):

$$\begin{aligned} \boldsymbol{T}_{hi}(t) = 2\exp\left(\lambda_i^{\mathrm{re}}t\right)\cdot\Big\{&\boldsymbol{V}_i^{\mathrm{re}}\left[\widetilde{T}_{0i}^{\mathrm{re}}\cdot\cos\left(\lambda_i^{\mathrm{im}}t\right) - \widetilde{T}_{0i}^{\mathrm{im}}\cdot\sin\left(\lambda_i^{\mathrm{im}}t\right)\right] \\ &-\boldsymbol{V}_i^{\mathrm{im}}\left[\widetilde{T}_{0i}^{\mathrm{re}}\cdot\sin\left(\lambda_i^{\mathrm{im}}t\right) + \widetilde{T}_{0i}^{\mathrm{im}}\cdot\cos\left(\lambda_i^{\mathrm{im}}t\right)\right]\Big\}. \end{aligned} \tag{20}$$

The amplitudes of harmonic variation of each of the real and the imaginary parts of the homogeneous solution are

$$A_{hi} = \sqrt{\left(\widetilde{T}_{0i}^{\mathrm{re}}\right)^2 + \left(\widetilde{T}_{0i}^{\mathrm{im}}\right)^2}. \tag{21}$$

- For the nonhomogeneous part of the solution (corresponding to the external load)

$$\begin{aligned} \boldsymbol{T}_{ni}(t) = 2\exp\left(\lambda_i^{\mathrm{re}}t\right)\cdot\frac{1}{\left(\lambda_i^{\mathrm{re}}\right)^2 + \left(\lambda_i^{\mathrm{im}}\right)^2}\cdot\Big\{&\boldsymbol{V}_i^{\mathrm{re}}\Big[\exp\left(-\lambda_i^{\mathrm{re}}t\right)\cdot\left(\lambda_i^{\mathrm{re}}\widetilde{F}_i^{\mathrm{re}} + \lambda_i^{i}\widetilde{F}_i^{i}\right) \\ &+\left(\lambda_i^{\mathrm{re}}\widetilde{F}_i^{\mathrm{im}} - \lambda_i^{\mathrm{im}}\widetilde{F}_i^{\mathrm{re}}\right)\cdot\sin\left(\lambda_i^{\mathrm{im}}t\right) - \left(\lambda_i^{\mathrm{re}}\widetilde{F}_i^{\mathrm{re}} + \lambda_i^{\mathrm{im}}\widetilde{F}_i^{\mathrm{im}}\right)\cdot\cos\left(\lambda_i^{\mathrm{im}}t\right)\Big] \\ &+\boldsymbol{V}_i^{\mathrm{im}}\Big[-\exp\left(-\lambda_i^{\mathrm{re}}t\right)\cdot\left(\lambda_i^{\mathrm{re}}\widetilde{F}_i^{\mathrm{im}} - \lambda_i^{\mathrm{im}}\widetilde{F}_i^{\mathrm{re}}\right) \\ &+\left(\lambda_i^{\mathrm{re}}\widetilde{F}_i^{\mathrm{re}} + \lambda_i^{\mathrm{im}}\widetilde{F}_i^{\mathrm{im}}\right)\cdot\sin\left(\lambda_i^{\mathrm{im}}t\right) + \left(\lambda_i^{\mathrm{re}}\widetilde{F}_i^{\mathrm{im}} - \lambda_i^{\mathrm{im}}\widetilde{F}_i^{\mathrm{re}}\right)\cdot\cos\left(\lambda_i^{\mathrm{im}}t\right)\Big]\Big\}. \end{aligned} \tag{22}$$

The amplitudes of harmonic variation of each of the real and the imaginary parts of the nonhomogeneous solution are

$$A_{ni} = \sqrt{\frac{\left(\widetilde{F}_i^{\mathrm{re}}\right)^2 + \left(\widetilde{F}_i^{\mathrm{im}}\right)^2}{\left(\lambda_i^{\mathrm{re}}\right)^2 + \left(\lambda_i^{\mathrm{im}}\right)^2}}. \tag{23}$$

3.4. Numerical methods used

An in-house computer program was written in Matlab (2002) to perform the transient modal analysis described above, with function *eig* used to solve the eigenvalue problem for system matrix $\boldsymbol{G}$. Matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ in Eq. (1) were determined indirectly, by solving a collection of trial problems using the commercial FE code Abaqus (2008).

4. Example solutions for friction clutches and brakes

4.1. Focal modes (in-plane sliding)

We consider a two-dimensional model of the sliding system shown in Fig. 1 with in-plane sliding, meaning that the model reflects modes of thermoelastic instability with variation in the sliding direction (focal hot spot pattern as defined by Anderson and Knapp (1990)), similar to that used by Lee and Barber (1993a) and in other studies (Zagrodzki et al., 2001; Zagrodzki and Truncone, 2003). It was demonstrated in these publications that the model is capable of realistic prediction of TEI in disk-type clutches and brakes. Although only two sliding interfaces are included in the current example, the model can easily be extended to a larger number. Solutions for the focal modes are harmonic in the sliding direction (Lee and Barber, 1993a). The horizontal length l_x in the model represents a single wavelength and therefore cyclic symmetry boundary conditions for the thermal and elastic problems are applied at the ends $x = 0$ and $x = l_x$. Stability of a specific mode depends on the wavelength and we are generally interested in those that show the greatest degree of instability. Since the wavelength of the most unstable mode is not known in advance, it is found by a series of trial simulations as in Zagrodzki and Truncone (2003). The model parameters used in calculations are representative of an automotive transmission wet clutch and the most important of them are listed in Table 1.

4.1.1. Mode shapes

In the examples presented, the dimension of the nodal temperature vector $\boldsymbol{T}$ is over 700 which is also the number of eigenmodes. It was found that there occur multiple eigenvalues but the modal matrix is non-singular or, in other words, all eigenmodes are linearly independent. Consequently, the transformation defined by Eq. (13) is possible. Selected mode shapes are shown in Fig. 2. The modes are normalized so that the Euclidean norm is 1. Mode numbers referenced are assigned arbitrarily, for convenience of this discussion. For complex modes, only the real part is plotted. (The imaginary part has the shape of the real part with a shift of $\pi/2$ in the sliding direction.) The linear scales in the x and y directions used in Fig. 2 are different, with an aspect ratio of around 10; different modes are shown with different view angles. Corresponding eigenvalues are listed in Table 2.

The modes are grouped in three patterns: focal antisymmetric, meaning modes with alternate locations of maximum temperatures at the two surfaces (Modes 1–3), focal symmetric (4–6) and non-focal (7–9). All focal modes are complex. Non-focal modes, with shapes not varying in the sliding direction, are real. The latter represent the background process for a perfectly uniform pressure distribution. None of them has a positive eigenvalue and hence all are stable. Mode 7, in particular, has zero eigenvalue and the shape characterized by constant value across the domain. It represents a uniform bulk temperature rise.

Table 1
Parameters and properties used in calculations.

Parameter	Value
Dimensional parameters	
Thickness of the layer of friction material (mm)	0.6
Thickness of steel disk (mm)	1.7
Friction material properties	
Modulus of elasticity (MPa)	75
Coefficient of thermal expansion (K^{-1})	6×10^{-5}
Conductivity (W/mK)	0.246
Specific heat (J/kg K)	1550
Density (kg/m^3)	733
Steel properties	
Modulus of elasticity (MPa)	2×10^5
Coefficient of thermal expansion (K^{-1})	1.2×10^{-5}
Conductivity (W/mK)	42
Specific heat (J/kg K)	452
Density (kg/m^3)	7800
Friction coefficient	0.13
Sliding speed (m/s)	37.5

Note that each focal mode, whether symmetric or antisymmetric, shows distinct temperature variation only across the stationary component while there is minor variation across the moving components, or vice versa. In other words, relative motion of the components is reflected in the mode shapes so that they have a distinct pattern in one of the components while the shape variation in the other, relatively moving component, is confined to a thin layer next to the interface. By contrast, many of non-focal modes show significant temperature variation across all the components.

It is known from previous TEI studies (Lee and Barber, 1993a; Yi et al., 2000; Zagrodzki, 1990) that antisymmetric modes are usually the dominant unstable modes in disk-type clutches and brakes with two (or more) friction surfaces. Indeed, antisymmetric Mode 1 has a positive real part of the eigenvalue $\mathrm{Re}(\lambda_1) = 13.60\ \mathrm{s}^{-1}$, and hence this mode is strongly unstable. The imaginary part of the eigenvalue represents the speed of migration of the mode shape in the sliding direction. The linear migration speed with respect to the coordinate system, which is fixed to the central steel plate (Fig. 1), is $c_i = \mathrm{Im}(\lambda_i) \cdot l_x/2\pi$. With $\mathrm{Im}(\lambda_1) = 29.82\ \mathrm{s}^{-1}$ and $l_x = 28.5 \times 10^{-3}$ m, the migration speed $c_1 = 0.135$ m/s. This speed is a very small fraction of the sliding speed $v_s = 37.5$ m/s. Hence, the mode is almost stationary with respect to the steel plate, which is a known phenomenon in sliding systems with thermally dissimilar materials (Lee and Barber, 1993a; Yi et al., 2000; Decuzzi et al., 2001). Nonetheless, the low-speed migration is reflected in the mode shape: in the contour plot of Mode 1 a skew bias of the temperature field reflecting phase difference across the thickness of the steel plate which is caused by migration is clearly visible. Mode 2, by contrast, moves at speed $c_2 = 37.50$ m/s and hence is stationary relative to the friction material. This mode, however, is stable with $\mathrm{Re}(\lambda_2) = -7.63\ \mathrm{s}^{-1}$. Antisymmetric Mode 3 shows strong temperature oscillations with reversals across the thickness of the layer of friction material.

All symmetric modes are stable in the example presented. The analysis, however, was performed for the length l_x corresponding to the most unstable antisymmetric mode and therefore the symmetric mode with the greatest real part was not captured. It was found that as the wavelength l_x increases, the real part of the eigenvalue of Mode 5 increases and it becomes dominant among the symmetric modes; this mode, nonetheless, is stable in the whole practically admissible range of wavelengths at the speed of 37.5 m/s. Mode 5, with temperature variation in the steel plate, has very small migration speed relative to it. Modes 4 and 6, on the other hand, move fast and are almost stationary relative to the friction material.

Interestingly, the antisymmetric Mode 2 has a counterpart in the symmetric Mode 4 with identical eigenvalue. Similarly, Modes 3 and 6 are other examples of multiple eigenvalues with linearly independent eigenvectors.

4.1.2. Nonhomogeneous solution

We assume a small geometric imperfection in the form of a sinusoidal waviness of the central layer representing the steel disk in the model shown in Fig. 1. The imperfection included in the elastic model produces some variation in the overall pressure expressed by Eq. (1). The pattern of pressure variation resulting from the waviness is obviously antisymmetric. The magnitude of the waviness, estimated in Zagrodzki and Truncone (2003) based on measurements of surface profiles of steel disks in an automotive automatic transmission clutch, was used in this study. The wavi-

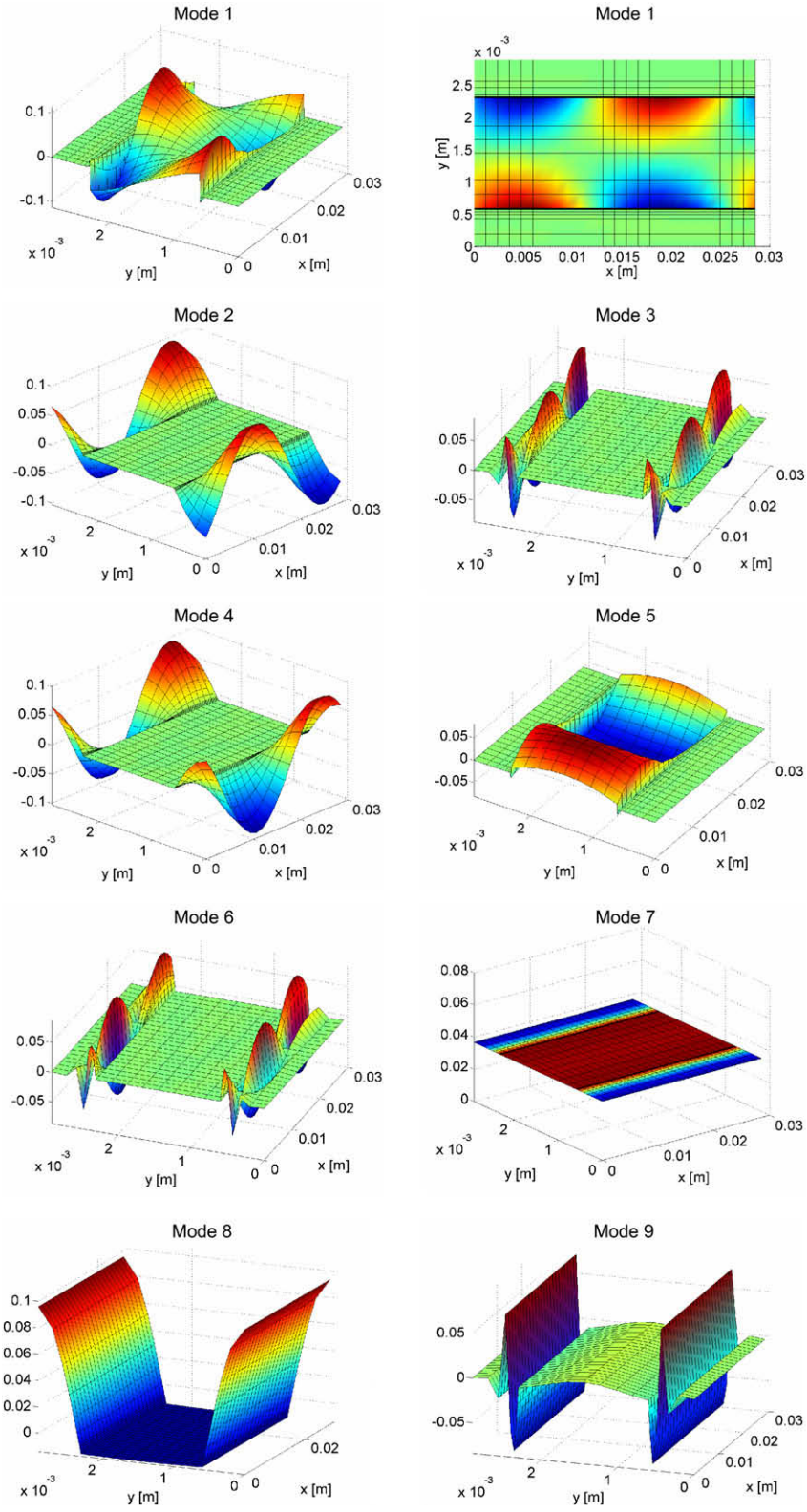

Fig. 2. Mode shapes, in-plane sliding.

ness corresponding to the wavelength l_x under consideration is generally small so that the resulting variation of contact pressure is a fraction, usually less than 1/10, of the nominal contact pressure. Table 2 shows the effect of pressure variation with amplitude of 0.08 MPa on the load expressed in modal variables. In the column representing the antisymmetric load pattern, non-zero load components occur in the rows corresponding to antisymmetric modes. As expected, there are no load contributors from this pattern of pressure variation to either focal symmetric modes or non-focal modes. A symmetric pattern of pressure distribution

Table 2
Problem with in-plane sliding: eigenvalues, modal loads and amplitudes of nonhomogeneous solution.

Mode shape and eigenvalue λ (s^{-1})				Load in modal variables $\tilde{F}$ (s^{-1}) and corresponding amplitude of nonhomogeneous solution A_n (–)								
Mode pattern	Mode number i	Re(λ_i)	Im(λ_i)	Pattern of pressure distribution due to geometry imperfection								
				Antisymmetric			Symmetric			Uniform		
				$\mathrm{Re}(\tilde{F}_i)$	$\mathrm{Im}(\tilde{F}_i)$	A_{ni}	$\mathrm{Re}(\tilde{F}_i)$	$\mathrm{Im}(\tilde{F}_i)$	A_{ni}	$\mathrm{Re}(\tilde{F}_i)$	$\mathrm{Im}(\tilde{F}_i)$	A_{ni}
Focal antisymmetric	1	13.60	29.82	−1127.38	−1039.90	46.80	0	0	0	0	0	0
	2	−7.63	8268.13	0.98	0.84	1.560e−4	0	0	0	0	0	0
	3	−115.89	8282.24	−22.02	−10.13	0.293e−2	0	0	0	0	0	0
Focal symmetric	4	−7.63	8268.12	0	0	0	−0.97	0.85	1.560e−4	0	0	0
	5	−9.07	8.39	0	0	0	−609.01	−480.16	62.79	0	0	0
	6	−115.89	8282.23	0	0	0	−18.48	−15.69	0.293e−2	0	0	0
Non-focal	7	0	0	0	0	0	0	0	0	−36155.6	0	36155.6
	8	−1.75	0	0	0	0	0	0	0	17958.4	0	10275.1
	9	−109.20	0	0	0	0	0	0	0	−26123.9	0	239.2

induced by geometry imperfection, attributed to thickness variation of the central layer, was examined as well and Table 2 shows the corresponding modal load. Only symmetric modes are excited by this form of pressure variation while it does not excite other modes. Further, a uniform pressure $p = 1$ MPa, corresponding to perfect geometry, contributes only to the non-focal modes that represent the ideal background process.

As discussed in Section 3.2, the nonhomogeneous part of the thermoelastic equation, which represents external loads, including the effects of possible contact pressure non-uniformity due to geometric imperfection, acts differently in terms of the excitation of unstable thermoelastic modes than the initial perturbation of the temperature field. There is no temperature variation associated with the former and there occurs a transitional phase during which it is being produced. Eq. (19) shows that the transitional component in the thermoelastic field induced by the nonhomogeneous part changes exponentially and it decays for an unstable mode. For example, for the unstable antisymmetric Mode 1 with Re (λ_1) = 13.60 s^{-1} it takes only 0.169 s for the transitional component to be reduced by factor of 10. In other words, the temperature field forms very quickly if a strongly unstable mode is excited by a compatible external load pattern; then the unstable mode develops like the one excited by the initial temperature variation. Another interesting issue is the relative strength of excitation of unstable modes by the initial condition and by the nonhomogeneous part. It can be characterized by the amplitudes of the homogeneous and nonhomogeneous parts of the solution given by Eqs. (21) and (23), respectively. To compare the amplitudes, we normalize the mode shape of interest so that it produces the same contact pressure variation as that caused by the nonhomogeneous part. In other words, if the amplitude of the contact pressure variation caused by geometric imperfection is p_{amp}, the mode shape $\boldsymbol{V}_i$ is scaled by a scalar factor κ such that amplitude of pressure $\boldsymbol{p} = \boldsymbol{A} \cdot (\kappa \boldsymbol{V}_i)$ is also p_{amp}. For the unstable antisymmetric Mode 1 we obtain $A_{h1} = 19.38\ s^{-1}$ and $A_{n1} = 46.80\ s^{-1}$, hence the excitation by the pressure variation is 2.4 times stronger than that by the initial temperature variation.

An example solution of a nonhomogeneous problem obtained for uniform pressure of 1 MPa and a superposed antisymmetric pressure variation (due to waviness) of $0.08 \times \cos(2\pi x/l_x)$ MPa is presented in Fig. 3. It shows a rapidly growing unstable mode, which migrates in the sliding direction, superposed on the solution corresponding to constant pressure. Fig. 4 shows the nonuniform components of the contact pressure distribution in the very early phase (up to 0.03 s) of the transient process. The contribution from the assumed geometric imperfection, which is constant and fixed in the frame of reference used, and the contribution from thermoelastic deformation, which grows and migrates are shown separately. The complete pressure distribution is a sum of these two contributions and the constant pressure. It can be seen that the amplitude of the thermoelastic pressure variation exceeds that caused by imperfection very quickly, shortly after 0.01 s of slip.

4.2. Band modes (out-of-plane sliding)

Focal and band modes are manifestation of the same phenomenon of thermoelastic instability, which was well demonstrated by Yi et al. (2000) who obtained both types of mode with the same three-dimensional model. There are, however, important differences in physical nature between the focal and band modes and it often justifies treating them separately. In the problem representing focal modes, there is in-plane sliding while in that with band modes the sliding is out-of-plane. Consequently, the latter, unlike the former, does not involve convective heat transfer. This, in turn, leads to a quite different pattern of heat flow in the friction material; with focal modes, oscillatory temperature variation in the direction normal to the surface occurs across a skin layer of the material, with very short wavelength (Yi et al., 2000; Zagrodzki et al., 2001), whereas there is no such an effect with band modes.

We consider an axisymmetric model for band modes with the geometry shown in Fig. 5. It is defined in cylindrical polar coordinates rz. The same general formulation of the thermal and elastic problems as in the previous part will be used. In the present model, however, $v_c = 0$ in Eq. (2) and the component with the convection matrix $v_c \boldsymbol{L_C}$ vanishes. In addition, terms characteristic of axisymmetric geometry are included in the present thermal and elastic models and the local sliding speed is defined as $v_s = r\omega_s$, with r being the radial coordinate and ω_s the angular sliding speed. Coordinates $r = r_i$ and $r = r_o$ define natural boundaries of the domain and, unlike in the model with in-plane sliding, there are no cyclic symmetry boundary conditions there.

Models for focal modes and band modes can be considered complementary, but there is an obvious difference in the degree of simplification of geometry between them: in the former, one of the spatial dimensions is simply neglected while there is no such simplification in the latter. Therefore any quantitative comparisons of results from the two models need to be treated cautiously.

4.2.1. Mode shapes

In the example presented below all material properties, thicknesses of layers of materials, friction coefficient, and sliding speed are the same as in the example for the problem with in-plane sliding. Fig. 6 shows selected mode shapes obtained for this model and Table 3 includes corresponding eigenvalues. The system matrix $\boldsymbol{G}$ is again non-symmetric. However, its eigenvalues are predominantly real; among several hundred eigenvalues only a few are complex with very small imaginary parts. This means that the modes are practically stationary (do not migrate along contact interface). This

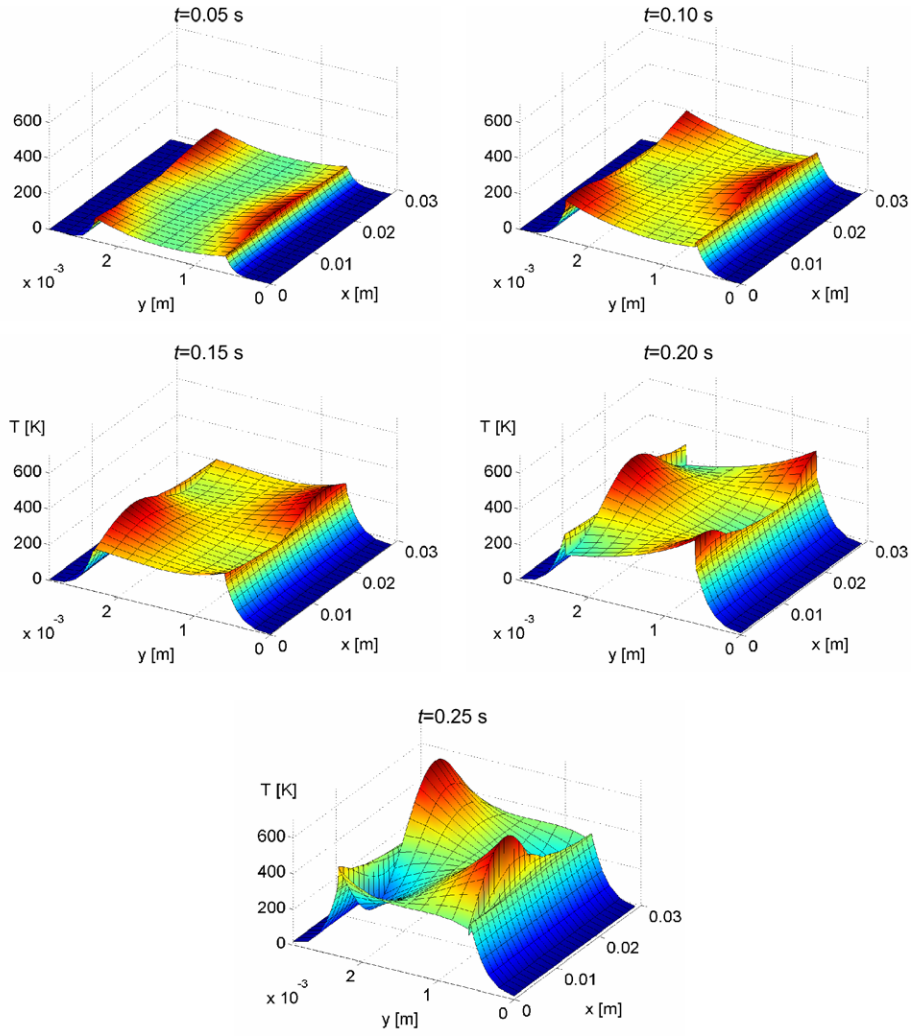

Fig. 3. Transient solution of the nonhomogeneous problem, in-plane sliding: temperature T (K).

effect is understandable: in the problem with in-plane sliding the migration is caused by the convective effect due to motion of layers of material while in the current problem there is no such mechanism. Mode 1 is antisymmetric. Close examination indicates that the mode is quasi-sinusoidal rather than sinusoidal since with the finite and relatively short radial length l_r the edges influence their shapes. The wavelength of Mode 1 is slightly less than l_r while that of Mode 2 is less than 2/3 of l_r. A contour plot of Mode 1 shows no skew bias of the temperature field in contrast to the migrating focal modes. Modes 2 and 3 are symmetric and quasi-sinusoidal. Interest-

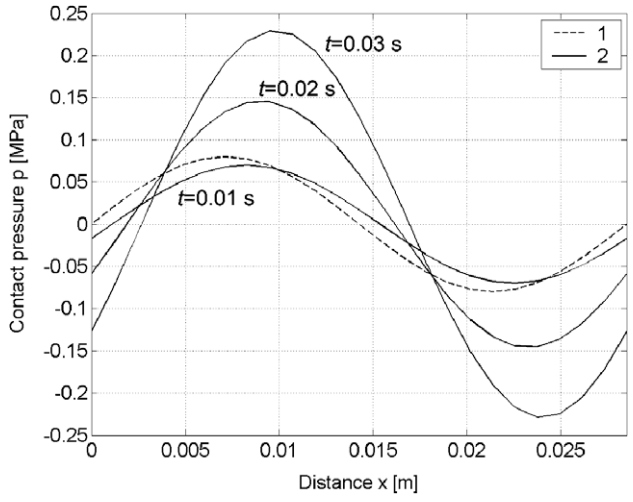

Fig. 4. Transient solution of the nonhomogeneous problem, in-plane sliding: evolution of the contact pressure p at the sliding interface. Pressure distribution (1) due to geometric imperfection and (2) due to thermal deformation.

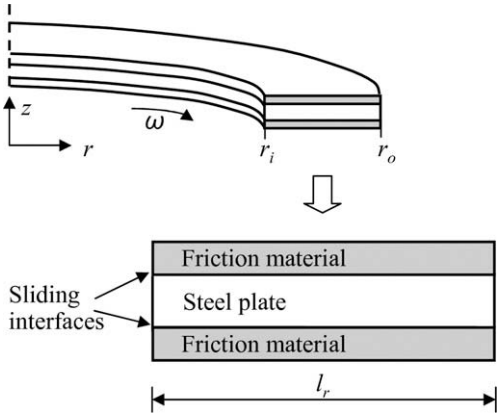

Fig. 5. Model geometry, out-of-plane sliding.

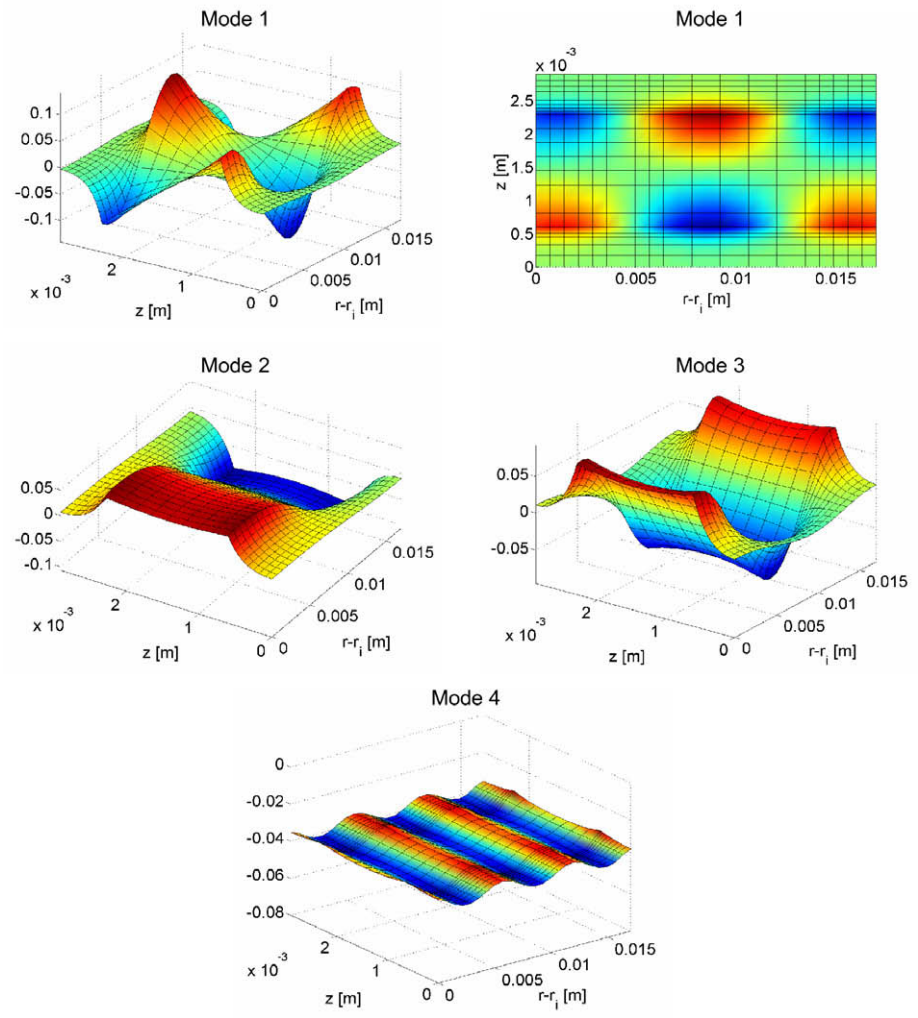

Fig. 6. Mode shapes, out-of-plane sliding.

ingly, all the band modes shown in Fig. 6 except one are unstable, while only one focal mode discussed in Section 4.1.1 was unstable. In this sense band modes tend to be less stable than the focal ones. This can be attributed to the lack of the convective term in band modes, which mechanism of heat transport is known to improve the stability (Lee and Barber, 1993a). On the other hand, the growth rate Re(λ) of the most unstable band mode (Mode 1) is lower than that of the unstable focal mode. It should be noted that the wavelength of the band modes is limited by radial length l_r of the friction surface, which is relatively small. Consequently, the modes with wavelengths greater than l_r that are anticipated to have higher degree of instability are not supported by this geometry.

4.2.2. Nonhomogeneous solution

Table 3 includes the eigenvalues of selected mode shapes and the corresponding modal loads. With real eigenvectors and eigenvalues, the modal loads are also real. As before, different load patterns, defined by the assumed contact pressure variation in isothermal state, were assumed. A strictly sinusoidal antisymmetric load pattern was used with wavelength the same as that of Mode 1 and amplitude of 0.05 MPa. As expected, this load pattern causes strong excitation of Mode 1. Similarly, for a symmetric load pattern a strictly sinusoidal pressure distribution corresponding to wavelength of Mode 3 was used and it can be observed that other symmetric modes, having different wavelengths, are also excited.

Table 3
Problem with out-of-plane sliding: eigenvalues, modal loads and amplitudes of nonhomogeneous solution.

Mode shape and eigenvalue λ (s^{-1})			Load in modal variables $\widetilde{F}$ (s^{-1}) and corresponding amplitude of nonhomogeneous solution A_n (–)							
Mode pattern	Mode number i	λ_i	Pattern of pressure distribution due to geometry imperfection							
			Antisymmetric		Symmetric		Uniform		Pseudo-uniform (with edge effect)	
			$\widetilde{F}_i$	A_{ni}	$\widetilde{F}_i$	A_{ni}	$\widetilde{F}_i$	A_{ni}	$\widetilde{F}_i$	A_{ni}
Band 'antisymmetric'	1	9.41	−1505.8	160.1	0	0	0	0	0	0
Band symmetric	2	5.46	0	0	153.4	28.1	1053.8	193.0	1044.0	191.2
	3	4.51	0	0	−787.4	174.6	409.0	90.7	446.5	99.0
	4	0	0	0	0	0	34640.5	34640.5	34640.5	34640.5

Excitation of the modes of different wavelength by a sinusoidal pressure variation, not seen with focal modes, can be attributed to two physical features of the current system. One of them is the effect of finite radial length l_r of the interface with boundaries at r_i and r_o and the other is the variation of sliding speed along the radius. Each of these features affects both the mode shapes and the modal loads.

The load pattern with assumed uniform isothermal pressure distribution excites, as expected, the mode with zero eigenvalue, representing a bulk temperature rise. But in addition a collection of symmetric modes are significantly excited. Mode 2, in particular, which represents monotonic temperature variation along radius, is clearly related to radial variation of sliding speed. The assumed uniform distribution of contact pressure in isothermal state is in fact artificial because in reality the pressure varies slightly in the vicinity of the edges. Therefore an additional load pattern, called pseudo-uniform, with this variation accounted for, was included. The difference in modal loads between this pattern and the uniform pressure for the modes shown is relatively small.

Fig. 7 presents a transient solution obtained with a small waviness of the steel plate causing antisymmetric pressure variation with amplitude of 0.05 MPa and uniform pressure of 1 MPa. It shows a rapid growth of the unstable Mode 1 along with mild growth of Mode 2 superimposed on the nominal background process represented mainly by Mode 4.

4.3. Out-of-plane sliding, single friction interface

An important observation from results shown in Table 3 is that unstable modes can be excited by the background process corresponding to a uniform pressure distribution. The excited modes are symmetric for the model geometry of Fig. 5. We have also studied the model shown in Fig. 8 with one sliding interface and the other non-sliding. This configuration is representative of steel plates at the end of the pack in a multidisk clutch. The end plate supporting the pack has a finite thickness and is locally supported in the axial direction. Fig. 9 shows selected mode shapes and Table 4 the eigenvalues and modal loads. The eigenvalues are again predominantly real. Mode shapes are neither symmetric nor antisymmetric. Two of the three modes shown in Fig. 9 are unstable and the uniform isothermal contact pressure produces substantial modal loads corresponding to these modes. Hence, the background process with uniform isothermal pressure excites non-symmetric unstable modes.

Fig. 10 shows an example of a transient solution with assumed uniform isothermal pressure distribution. The solution is dominated by very rapid growth of the unstable Mode 1 leading to early contact separation at around 0.1 s. This pattern of hot spots is well known from industrial practice (e.g. Zagrodzki and Farris, 1999). It has an obvious physical interpretation: a ring plate with heat uniformly applied to one surface deforms spherically which shape has

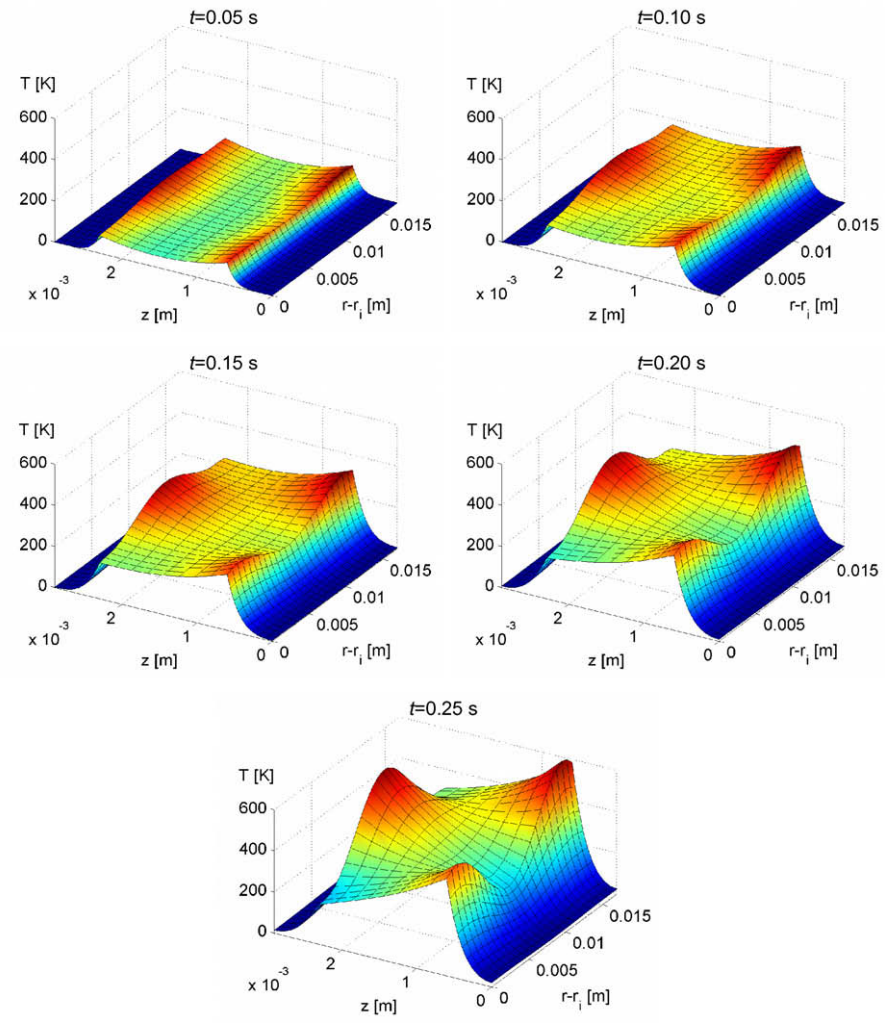

Fig. 7. Transient solution of the nonhomogeneous problem, out-of-plane sliding: temperature T (K).

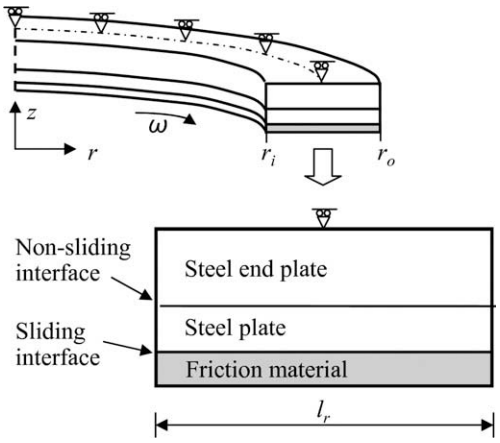

Fig. 8. Model geometry, out-of-plane sliding, single sliding interface.

a strong content corresponding to the unstable Mode 1. This means that the background process strongly excites this mode.

In the model for focal hot spots from Fig. 1, one of the sliding interfaces was also deactivated so that there was only one active interface. In this case modal loads corresponding to all antisymmetric and symmetric modes, including the unstable one, are zero with uniform pressure. This is because in a system with cyclic symmetry boundary conditions, the background process with uniform load produces uniform temperature/deformation along the interfaces which pattern is orthogonal with respect to these modes. This implies that the background process does not excite unstable modes in this case.

5. Conclusions

The formulation of the thermoelastic problem in friction clutches and brakes presented in the paper provides an evaluation of the stability of the system and a complete transient solution by modal superposition. The transient solution includes both the homogeneous part, corresponding to the initial condition, and the nonhomogeneous part, representing the background process. While it is well known that an initial temperature distribution that is consistent with an unstable mode grows exponentially, the behavior of the system in the absence of initial temperature variation was less understood. It becomes obvious from this discussion that the latter problem is of primary conceptual and practical significance. This study clearly exhibits excitation of unstable modes by the nonhomogeneous term. There is a transitional phase during which the temperature distribution consistent with an unstable mode is formed. The stronger the instability of the mode, the shorter is the duration of this phase. A sample quantitative evaluation for an automotive transmission multidisk wet clutch shows that the background process is often the primary trigger of unstable behavior that can lead to severe hot spots.

Two contributors are distinguished in the background process. One is the nominal process corresponding to uniform pressure distribution in isothermal state and the other is pressure variation caused by geometric imperfection or any other possible factors like design features. It was demonstrated that with a model configuration representative of so-called focal modes, the background process does not excite unstable modes unless pressure variation is involved. It was also shown that relatively small pressure variation, attributed to realistic imperfections, is sufficient to significantly excite the unstable mode within a short time interval. With other system configurations, like that representative of band modes, the background process may excite unstable modes even with perfectly uniform isothermal pressure distribution. In other words, the background process with an ideal pressure distribution produces deformation that has a strong modal content of unstable mode(s).

The model used for focal modes, involving in-plane sliding, includes a convective term representing the motion of solid components of the system relative to the frame of reference. It reflects a comprehensive class of sliding systems and can be extended to 3-D geometry without any new conceptual problems. However,

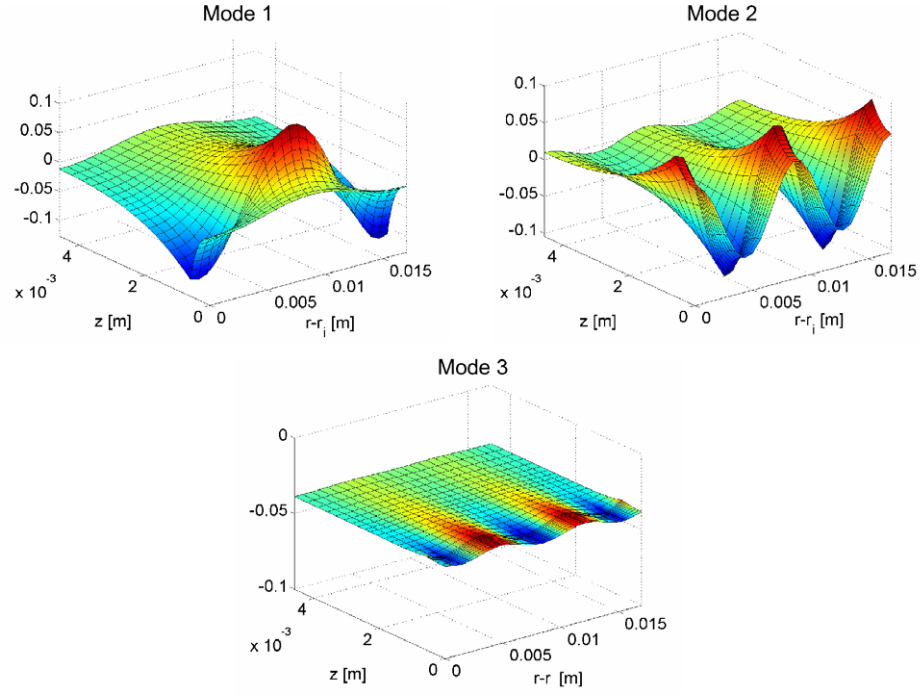

Fig. 9. Mode shapes, out-of-plane sliding, single sliding interface.

Table 4
Problem with out-of-plane sliding, single sliding interface: eigenvalues, modal loads and amplitudes of nonhomogeneous solution.

Mode shape and eigenvalue λ (s^{-1})			Load in modal variables $\widetilde{F}$ (s^{-1}) and corresponding amplitude of nonhomogeneous solution A_n (–)	
Mode pattern	Mode number i	λ_i	Pattern of pressure distribution: uniform	
			$\widetilde{F}_i$	A_{ni}
Band	1	4.4232	−8588.5	1941.7
	2	0.5696	2083.8	3658.4
	3	0	8639.0	8639.0

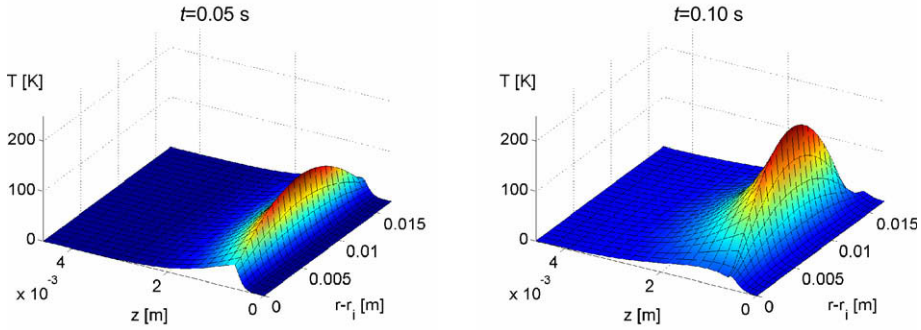

Fig. 10. Transient solution of the nonhomogeneous problem, out-of-plane sliding, single sliding interface: temperature T (K).

the resulting large-size non-symmetric eigenvalue problem is still a strong prohibitive factor in practical implementation. The extremely fine finite element mesh required to capture short-wave temperature variation in the skin of friction material is one of the major contributors to the computational intensity of this problem.

Appendix A. Finite element discretization of elastic problem with multiple sliding interfaces

Sliding contact conditions, without sticking or separation, are considered. The problem is linear and after finite element discretization it can be represented by a matrix equation of static equilibrium

$$\boldsymbol{K}\boldsymbol{u} = \boldsymbol{F}, \tag{A1}$$

where $\boldsymbol{u}$ is the vector of nodal displacements, $\boldsymbol{K}$ is the stiffness matrix, and $\boldsymbol{F}$ is the load vector. Eq. (A1) represents not only an ideal system with conforming surfaces but also the one with non-conforming surfaces under the condition that the load $\boldsymbol{F}$ enforces full contact. Consequently, it can represent a system with surface imperfections whenever sufficient preload is applied so that the contact region is fixed. The pressure variation at the sliding interfaces produced under preload has to be superimposed on the pressure determined below. Two contributions to the load vector $\boldsymbol{F}$ will be considered: one representing the external forces applied to the system $\boldsymbol{F}_a$ and the other representing the nodal loads corresponding to thermal strains $\boldsymbol{F}_t$. The latter can be expressed as

$$\boldsymbol{F}_t = \boldsymbol{\Psi}\boldsymbol{T}, \tag{A2}$$

where $\boldsymbol{\Psi}$ is a matrix and $\boldsymbol{T}$ is the vector of nodal temperatures. In the elastic problem the temperature is treated as a prescribed field variable. Hence, Eq. (A1) takes the form

$$\boldsymbol{K}\boldsymbol{u} = \boldsymbol{F}_a + \boldsymbol{\Psi}\boldsymbol{T} \tag{A3}$$

with solution in symbolic notation

$$\boldsymbol{u} = \boldsymbol{K}^{-1}(\boldsymbol{F}_a + \boldsymbol{\Psi}\boldsymbol{T}). \tag{A4}$$

We now consider a single jth component of the sliding system, e.g. single friction plate. The equation of static equilibrium of the component is

$$\boldsymbol{K}^j\boldsymbol{u}^j = \boldsymbol{F}^j, \tag{A5}$$

where $\boldsymbol{K}^j$, $\boldsymbol{u}^j$ and $\boldsymbol{F}^j$ are, respectively, the stiffness matrix, the displacement vector and the load vector of the component. $\boldsymbol{u}^j$ is a subset of $\boldsymbol{u}$ which can be expressed as

$$\boldsymbol{u}^j = \boldsymbol{Z}^j\boldsymbol{u}, \tag{A6}$$

where $\boldsymbol{Z}^j$ is an appropriate binary matrix. The load vector can be represented analogously to that for the whole system

$$\boldsymbol{F}^j = \boldsymbol{F}^j_{ap} + \boldsymbol{\Psi}^j\boldsymbol{T}^j, \tag{A7}$$

where $\boldsymbol{\Psi}^j$ is an appropriate matrix and the vector of nodal temperatures $\boldsymbol{T}^j$ is a subset of the vector $\boldsymbol{T}$,

$$\boldsymbol{T}^j = \boldsymbol{Z}^j_T\boldsymbol{T}, \tag{A8}$$

where $\boldsymbol{Z}^j_T$ is a binary matrix. Vector $\boldsymbol{F}^j_{ap}$ is more than a subset of $\boldsymbol{F}_a$ as it also includes forces representing contact pressure at the interfaces with adjacent components (which forces cancel in the global equation (A3) as internal ones). We can write

$$\boldsymbol{F}^j_{ap} = \begin{Bmatrix} \boldsymbol{P}^j \\ \boldsymbol{F}^j_a \end{Bmatrix}, \tag{A9}$$

where $\boldsymbol{P}^j$ is the vector of nodal forces representing interfacial pressure and $\boldsymbol{F}^j_a$ is a subset of $\boldsymbol{F}_a$ corresponding to the jth component. We can divide the vector of nodal displacements analogously to the vector $\boldsymbol{F}^j_{ap}$. As a result, two subsets of displacements are distinguished for the jth component: $\boldsymbol{u}^j_1$, which includes displacements of contact nodes normal to the interface, and $\boldsymbol{u}^j_2$, containing all the remaining displacements of the component. Eq. (A5) can now be rewritten as

$$\begin{bmatrix} \boldsymbol{K}^j_{11} & \boldsymbol{K}^j_{12} \\ \boldsymbol{K}^j_{21} & \boldsymbol{K}^j_{22} \end{bmatrix} \begin{Bmatrix} \boldsymbol{u}^j_1 \\ \boldsymbol{u}^j_2 \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{P}^j \\ \boldsymbol{F}^j_a \end{Bmatrix} + \begin{bmatrix} \boldsymbol{\Psi}^j_1 \\ \boldsymbol{\Psi}^j_2 \end{bmatrix} \boldsymbol{T}^j, \tag{A10}$$

where $\boldsymbol{K}^j_{\alpha\beta}$, α,β=1,2, are appropriate sub-matrices of $\boldsymbol{K}^j$ and $\boldsymbol{\Psi}^j_\alpha$ sub-matrices of $\boldsymbol{\Psi}^j$.

From Eq. (A10) we obtain

$$\boldsymbol{P}^j = -\boldsymbol{\Psi}^j_1 \boldsymbol{T}^j + \begin{bmatrix} \boldsymbol{K}^j_{11} & \boldsymbol{K}^j_{12} \end{bmatrix} \begin{Bmatrix} \boldsymbol{u}^j_1 \\ \boldsymbol{u}^j_2 \end{Bmatrix}. \tag{A11}$$

Substitution of solution (A4) and relations (A6) and (A8) into Eq. (A11) leads to

$$\boldsymbol{P}^j = \boldsymbol{A}^j \boldsymbol{T} + \boldsymbol{B}^j \boldsymbol{F}_a, \tag{A12}$$

where

$$\boldsymbol{A}^j = -\boldsymbol{\Psi}^j_1 \boldsymbol{Z}^j_T + \begin{bmatrix} \boldsymbol{K}^j_{11} & \boldsymbol{K}^j_{12} \end{bmatrix} \boldsymbol{Z}^j \boldsymbol{K}^{-1} \boldsymbol{\Psi} \quad \text{and} \quad \boldsymbol{B}^j = \begin{bmatrix} \boldsymbol{K}^j_{11} & \boldsymbol{K}^j_{12} \end{bmatrix} \boldsymbol{Z}^j \boldsymbol{K}^{-1}. \tag{A13}$$

We define the vector $\boldsymbol{P}$ which includes the nodal forces representing contact pressure at all interfaces of the sliding system. Vectors $\boldsymbol{P}^j$ are subsets of $\boldsymbol{P}$. Note that the vector of contact forces at any interface can be calculated from either of the two components that create the interface and we use only one of them. We repeat Eq. (A12) for a subset of the components of the system, which have to be selected in such a way that the vector $\boldsymbol{P}$ of interfacial forces is complete. This set of matrix equations (A12) can be collected in a single matrix equation of the form

$$\boldsymbol{P} = \boldsymbol{A}\boldsymbol{T} + \boldsymbol{B}\boldsymbol{F}_a, \tag{A14}$$

where $\boldsymbol{A}$ is a block matrix created by matrices $\boldsymbol{A}^j$ from this set and, similarly, $\boldsymbol{B}$ is created by matrices $\boldsymbol{B}^j$. The dimension of vector $\boldsymbol{P}$ is implied by context and when required this vector is expanded (scattered) to the dimension of the vector $\boldsymbol{T}$ and matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are expanded accordingly.

Appendix B. Finite element discretization of conductive-convective heat transfer

We consider the two-dimensional transient heat transfer problem in a sliding system. Using Eulerian coordinates the problem in the jth component of the system is described by the equation

$$c^j \rho^j \left(\frac{\partial T^j}{\partial t} + v^j_c \frac{\partial T^j}{\partial x} \right) - K^j \nabla^2 T^j = 0 \text{ in } \Omega^j, \tag{B1}$$

where $T^j = T^j(t,x,y)$ is temperature; t is time; v^j_c is velocity in x direction; c^j, ρ^j and K^j are specific heat, mass density, and thermal conductivity, respectively, and Ω^j is the domain of the component. Eq. (B1) includes a convective term which reflects the mass flow of the jth component with respect to the coordinate system. (For components that are stationary with respect to the coordinates $v^j_c = 0$ and the convective term vanishes. Also, v^j_c is uniform within a component which fact will be utilized in the following derivations.)

The boundary and initial conditions for Eq. (B1) have the form

$$-K^j \frac{\partial T^j}{\partial n} = q^j \text{ on } \partial\Omega^j \quad \text{and} \quad T^j(0,x,y) = T^j_0(x,y), \tag{B2}$$

where q^j is the heat flux applied at the boundary $\partial\Omega^j$ and n is the outward normal to the boundary.

The weak form of Eq. (B1) is

$$\int_{\Omega^j} W \left[c^j \rho^j \left(\frac{\partial T^j}{\partial t} + v^j_c \frac{\partial T^j}{\partial x} \right) - K^j \nabla^2 T^j \right] dx\,dy = 0, \tag{B3}$$

where W is a weighting function. On the use of Green's theorem Eq. (B3) can be rewritten as

$$\int_{\Omega^j} W c^j \rho^j \frac{\partial T^j}{\partial t}\,dx\,dy + \int_{\Omega^j} W v^j_c \frac{\partial T^j}{\partial x}\,dx\,dy + \int_{\Omega^j} \nabla W \cdot K^j \nabla T^j\,dx\,dy - \int_{\partial\Omega^j} W K^j \frac{\partial T^j}{\partial n}\,ds = 0 \tag{B4}$$

Spatial discretization of Eq. (B4) consists in piecewise approximation of temperature $T^j(t,x,y) = \boldsymbol{N}(x,y)\boldsymbol{T}^j(t)$, where $\boldsymbol{N}$ is a row vector containing shape functions $N_i(x,y)$ and $\boldsymbol{T}^j$ is a column vector of nodal temperatures $T^j_i(t)$, $i = 1,\ldots,n^j$, with n^j denoting number of nodes in the finite element mesh over jth component. In the Galerkin finite element method the weighting functions W are the same as the shape functions N. However, the Galerkin formulation is known to lead to a numerically unstable algorithm if the convective term in Eq. (B1) is substantial; more specifically, the method is only suitable for low mesh Peclet numbers $\gamma \leqslant 2$, where $\gamma = v_c h_x / K$ and h_x is the linear dimension of the element in the direction of mass flow. In sliding systems such as friction clutches and brakes, speeds are generally high and the problem is strongly convection dominated with γ possibly as high as 10^4. A common finite element approach for solving these types of problems employs the upwind formulation which is realized by using appropriately designed non-symmetric weighting functions W (Brooks and Hughes, 1982), different from the shape functions N. This formulation is known as Petrov–Galerkin and it was used in this study. Upon integration of Eq. (B4) we obtain

$$\boldsymbol{C}^j \dot{\boldsymbol{T}}^j + \boldsymbol{L}^j_{\boldsymbol{D}} \boldsymbol{T}^j + v^j_c \boldsymbol{L}^j_{\boldsymbol{C}} \boldsymbol{T}^j = \boldsymbol{Q}^j, \tag{B5}$$

where

$$\boldsymbol{C}^j = \int_{\Omega^j} c^j \rho^j \boldsymbol{W}' \boldsymbol{N}\,dx\,dy \tag{B6}$$

is the thermal capacity matrix;

$$\boldsymbol{L}^j_{\boldsymbol{D}} = \int_{\Omega^j} K^j \nabla \boldsymbol{W}' \nabla \boldsymbol{N}\,dx\,dy \tag{B7}$$

is the conductivity matrix;

$$\boldsymbol{L}^j_{\boldsymbol{C}} = \int_{\Omega^j} \boldsymbol{W}' \frac{\partial \boldsymbol{N}}{\partial x}\,dx\,dy \tag{B8}$$

and the product $v^j_c \boldsymbol{L}^j_{\boldsymbol{C}}$ is the convection matrix; and

$$\boldsymbol{Q}^j = \int_{\partial\Omega^j} \boldsymbol{W}' K^j \frac{\partial T^j}{\partial n}\,ds \tag{B9}$$

is the thermal load vector. $\boldsymbol{W}$ is a row vector containing weighting functions W_i and $\boldsymbol{W}'$ denotes its transpose. The convection matrix $v^j_c \boldsymbol{L}^j_{\boldsymbol{C}}$ is non-symmetric by the physical nature of the process; with non-symmetric weighting functions, also the thermal capacitance matrix $\boldsymbol{C}^j$ is non-symmetric (Brooks and Hughes, 1982), but the matrix $\boldsymbol{L}^j_{\boldsymbol{D}}$ is symmetric.

Note that temperature gradient $\partial T^j/\partial n$ in Eq. (B9) at the part of the boundary $\partial\Omega^j$ of jth component which constitutes an interface $\partial\Omega^{jk}$ with adjacent kth component is unknown. However, the energy conservation condition at the interface $\partial\Omega^{jk}$ furnishes the following equation coupling the temperature fields in the two components

$$K^j \frac{\partial T^j}{\partial n}\bigg|_{\partial\Omega^{jk}} - K^k \frac{\partial T^k}{\partial n}\bigg|_{\partial\Omega^{jk}} = q^{jk}, \tag{B10}$$

where q^{jk} is the frictional heat flux generated at the interface. This equation along with temperature continuity at the interface

$$T^j\big|_{\partial\Omega^{jk}} = T^k\big|_{\partial\Omega^{jk}} \tag{B11}$$

uniquely defines the interfacial boundary conditions. Eq. (B5) for all components can now be collected in a single matrix equation repre-

senting the whole sliding system, in which process the temperature gradients present in Eq. (B9) cancel between components upon using Eqs. (B10) and (B11), and we obtain

$$\boldsymbol{C}\dot{\boldsymbol{T}} + \boldsymbol{L_D}\boldsymbol{T} + v_c\boldsymbol{L_C}\boldsymbol{T} = \boldsymbol{Q} \tag{B12}$$

with the thermal load vector defined as

$$\boldsymbol{Q} = \int_{\partial\Omega^{\text{intfce}}} \boldsymbol{W}\boldsymbol{q}\, ds, \tag{B13}$$

where $\partial\Omega^{\text{intfce}}$ represents all interfaces in the system and $\boldsymbol{q}$ is the vector of frictional heat flux applied at these interfaces.

References

The Abaqus Software, 2008. Analysis User's Manual, Version 6.8, Dassault Systèmes Simulia Corp., Providence, RI, USA.

Afferrante, L., Ciavarella, M., Decuzzi, P., Demelio, G., 2003. Transient analysis of frictionally excited thermoelastic instability in multi-disk clutches and brakes. Wear 254, 136–146.

Al-Shabibi, A.M., Barber, J.R., 2002. Transient solution of a thermoelastic instability problem using a reduced order model. Int. J. Mech. Sci. 44, 451–464.

Anderson, A.E., Knapp, R.A., 1990. Hot spotting in automotive friction systems. Wear 135, 319–337.

Barber, J.R., 1969. Thermoelastic instabilities in the sliding of conforming solids. Proc. R. Soc. Lond. A 312, 381–394.

Brooks, A.N., Hughes, T.J.R., 1982. Streamline upwind/Petrov–Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations. Comput. Meth. Appl. Mech. Eng. 32, 199–259.

Choi, J.H., Lee, I., 2004. Finite element analysis of transient thermoelastic behaviors in disk brakes. Wear 257, 47–58.

Coddington, E.A., Carlson, R., 1997. Linear Ordinary Differential Equations. Society for Industrial and Applied Mathematics, Philadelphia.

Decuzzi, P., Ciavarella, M., Monno, G., 2001. Frictionally excited thermoelastic instability in multi-disk clutches and brakes. ASME J. Tribol. 123, 865–871.

Deif, A.S., 1982. Advanced Matrix Theory for Scientists and Engineers. Halsted Press Division, Wiley, New York.

Dow, T.A., Burton, R.A., 1972. Thermoelastic instability of sliding contact in the absence of wear. Wear 19, 315–328.

Du, S., Zagrodzki, P., Barber, J.R., Hulbert, G.M., 1997. Finite element analysis of frictionally excited thermoelastic instability. J. Therm. Stress. 20, 185–201.

Kao, T.K., Richmond, J.W., Douarre, A., 2000. Brake disc hot spotting and thermal judder: an experimental and finite element study. Int. J. Vehicle Des. 23 (3/4), 276–296.

Krempaszky, C., Lippmann, H., 2005. Frictionally excited thermoelastic instability of annular plates under thermal pre-stress. ASME J. Tribol. 127, 756–765.

Lee, K., Barber, J.R., 1993a. Frictionally excited thermoelastic instability in automotive disk brakes. ASME J. Tribol. 115, 607–614.

Lee, K., Barber, J.R., 1993b. The effect of shear tractions on frictionally excited thermoelastic instability. Wear 160, 237–242.

Lee, K., Brooks Jr., F.W., 2003. Hot spotting and judder phenomena in aluminum drum brakes. ASME J. Tribol. 125, 44–51.

Li, J., Barber, J.R., 2004. Transient solution of the inhomogeneous thermoelastic contact problem using fast speed expansion. In: IJTC04, 2004 ASME/STLE Joint Tribology Conference, Long Beach, CA, USA.

Li, J., Barber, J.R., 2008. Solution of transient thermoelastic contact problems by fast speed expansion method. Wear 265, 402–410.

Matlab, 2002. Version 6.5, The MathWorks Inc., Natick, MA, USA.

Voldrich, J., 2007. Frictionally excited thermoelastic instability in disc brakes – transient problem in the full contact regime. Int. J. Mech. Sci. 49, 129–137.

Yi, Y.B., Barber, J.R., Zagrodzki, P., 2000. Eigenvalue solution of thermoelastic instability problems using Fourier reduction. Proc. R. Soc. Lond. A, 456, 2799–2821.

Zagrodzki, P., 1990. Analysis of thermomechanical phenomena in multidisk clutches and brakes. Wear 140, 291–308.

Zagrodzki, P., 2003. Modal analysis of the general case of a sliding system with frictionally excited thermoelastic instability. In: Society of Engineering Sciences Meeting, Ann Arbor, MI, USA.

Zagrodzki, P., Farris, T.D., 1999. Analysis of temperatures and stresses in wet friction disks involving thermally induced changes of contact pressure. SAE Technical Paper 982035, SAE Transactions, 107, Sect. 2, pp. 360–367.

Zagrodzki, P., Truncone, S.A., 2003. Generation of hot spots in a wet multidisk clutch during short-term engagement. Wear 254, 474–491.

Zagrodzki, P., Zhao, W., 2008. Thermoelastic behaviour of an automatic transmission clutch with a 'finger' piston involving thermoelastic instability. Int. J. Vehicle Des. 48, 97–113.

Zagrodzki, P., Lam, K.B., Al Bahkali, E., Barber, J.R., 2001. Non-linear transient behavior of a sliding system with frictionally excited thermoelastic instability. ASME J. Tribol. 123, 699–708.